



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Estimation in the Inverse Gaussian Regression Model

Ravinder Kaur Singh

A Thesis
in
The Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montréal, Québec, Canada

June 1992

©Ravinder Kaur Singh, 1992



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-80978-7

Canada

Abstract

Estimation in the Inverse Gaussian Regression Model

Ravinder Kaur Singh

This thesis provides a short review of methods of estimation of parameters for the Linear Model and resampling inference procedures using the inverse Gaussian distribution. In this particular case, we consider the jackknife procedure and bootstrap method with the inverse Gaussian Regression Model.

Chapter 1 provides a brief description of a *linear model*, the various methods of estimation of parameters in such a model and describes some of the important properties of the inverse Gaussian distribution.

Chapter 2 briefly sketches two standard methods of resampling inference, namely the jackknife and the bootstrap methods. We also review briefly in this chapter the use of jackknife and bootstrap in a standard linear model.

Chapter 3 considers the inverse Gaussian regression model and reviews the related literature. We consider the adaptation of the resampling methods for this model.

Chapter 4 presents a numerical study of the techniques developed in chapter 3.

Acknowledgements

I am grateful to my advisor Prof. Yogendra Chaubey for accepting me as a student, and guiding me in my thesis. He has always been encouraging, patient, and available for discussion.

I am thankful to Prof. J. Garido and Prof. T.N. Stivastava for agreeing to be on my thesis committee and for helpful comments which have improved the writing of thesis.

Contents

1 Some Preliminaries of the Linear Model and the Inverse Gaussian Distribution	1
1.1 Introduction	1
1.2 General Linear Models	2
1.3 Estimation of the Regression Parameters.....	3
1.3.1 Method of Least Squares	4
1.3.2 Method of Weighted Least Squares	4
1.3.3 Method of Maximum Likelihood.....	6
1.4 The Inverse Gaussian Distribution and its Properties	7
2 Resampling Inference in a Standard Linear Model	10
2.1 Introduction	10
2.2 Jackknife and Bootstrap for a Single parameter.....	11
2.2.1 The Jackknife Method	11
2.2.2 The Bootstrap method	12
2.3 Jackknife and Bootstrap for the Regression Parameter	15
2.3.1 The Jackknife Method	15
2.3.2 The Bootstrap Method	17
3 Resampling Inference in an Inverse Gaussian Regression Model	22
3.1 Introduction	22
3.1.1 The Model	23
3.1.2 Maximum Likelihood Estimation	25
3.2 The Jackknife Method	27
3.3 The Bootstrap Method	29
4 A Numerical Study of Various Estimators	32

4.1 Design of the Simulation Study	32
4.2 Conclusions	34
4.2.1 The Bias of Estimators	34
4.2.2 The MSE's of the Estimators	47
4.2.3 Estimator of the MSE of MLE	61
Bibliography	68
Appendix	70

Chapter 1

SOME PRELIMINARIES OF THE LINEAR MODEL AND THE INVERSE GAUSSIAN DISTRIBUTION

1.1 Introduction

Scientists employ mathematical models to express the relationship between various variables of interest. For example, the economist relates the price of a commodity to demand and supply, the physicist expresses the interplay between the force of a moving body and its mass and the velocity, and the chemist relates the solubility of a solid in a liquid to environmental conditions as well as the characteristics of the solid and the liquid. Or agriculturists could be interested in the relationship between the number of apples in a tree in an orchard and the amount of fertilizers the tree received. All such relations can be formulated in terms of mathematical equations relating the variables involved, which form a model of reality in much the same way as a photographer attempts to capture an image and, hence, model a subject.

The mathematical model may be linear or non linear. In this thesis we are

interested in the linear model where we assume a linear relationship between different variables, as described in the next section. Section 1.3 gives various methods of estimation of parameters for such a model. We will also be concerned here with the inverse Gaussian regression model : some of its properties are summarized in section 1.4.

1.2 General Linear Models

The general linear statistical model relating the response variable Y to a set of $(p-1)$ explanatory variables X_1, X_2, \dots, X_{p-1} is written as

$$Y = \sum_{i=0}^{p-1} X_i \beta_i + \mathcal{E} \quad (1.1)$$

where $X_0 = 1$ and $\beta_0, \beta_1, \dots, \beta_{p-1}$ are unknown parameters and \mathcal{E} is the hypothetical error such that

$$E(\mathcal{E}) = 0, \quad \text{Cor}(\mathcal{E}) = \sigma^2 I. \quad (1.2)$$

Given a series of n observations on Y , denoted by Y_1, Y_2, \dots, Y_n , Y_i is supposed to have been observed for $(X_{i1}, X_{i2}, \dots, X_{ip-1})$ as the values of $(X_1, X_2, \dots, X_{p-1})$ respectively. We write n equations derived from (1.1) as

$$Y_{n \times 1} = X_{n \times p} \beta_{p \times 1} + \mathcal{E}_{n \times 1} \quad (1.3)$$

where

$$Y_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X_{n \times p} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np-1} \end{bmatrix}$$

$$\beta_{p \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \text{ and } \mathcal{X}_{n \times 1} = \begin{bmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \\ \vdots \\ \mathcal{X}_n \end{bmatrix}$$

and further we have

$$E(\mathcal{X}_t) = 0, \quad V(\mathcal{X}_t) = \sigma^2 I$$

and

$$Cov(\mathcal{X}_t, \mathcal{X}_l) = 0, \quad \forall \quad t \neq l \quad t = 1, 2, \dots, n \quad (1.4)$$

In some cases the covariance matrix may not be of the form as in (1.4) and the identity matrix I may be substituted by a symmetric positive definite matrix denoted by Σ .

Further assumptions about the distributional properties of \mathcal{X} or Y , or further assumptions about the structure of V can be stated in many cases as part of the model. In the next section we present basic methods of estimating the parameter vector β .

1.3 Estimation of the Regression Parameters

We will discuss three methods to estimate the unknown parameter β , namely

- (i) The method of Least Squares,
- (ii) The method of Weighted Least Squares,
- (iii) The Maximum Likelihood Method.

These are described below

1.3.1 Method of Least Squares.

The method of least squares involves minimizing the sum of squares of deviations of the observations from their expectations, namely,

$$Q = \sum_{i=1}^n (Y_i - \sum_{j=0}^{p-1} X_{ij} \beta_j)^2 \quad (1.5)$$

with respect to $(\text{vector}) \beta$, which can be written as

$$Q(\beta) = (Y - X\beta)'(Y - X\beta).$$

Assuming that X is of full rank, the solution of the minimization is easily obtained as

$$\beta = (X'X)^{-1}X'Y. \quad (1.6)$$

It is a well known fact that least squares estimation does not pre-suppose any distributional properties of the \mathcal{E}_i 's other than finite (in our case zero) mean and equal variances.

Sometimes it so happens that all the observations do not have equal variance. In that case, the least squares method is not appropriate. We can resort to the weighted least squares method of estimation in that situation. This method is briefly described below.

1.3.2 Method of Weighted Least Squares

For weighted least squares, each element of \mathcal{E} is uncorrelated with each other

but the variance need not be constant i.e.

$$V(\mathcal{E}) = \sigma^2 \Sigma \quad (1.7)$$

where Σ is the matrix

$$\Sigma = W^{-1} = \begin{bmatrix} \frac{1}{w_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{w_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{w_n} \end{bmatrix}$$

and σ^2 is generally an unknown parameter, but the $w_i > 0$ are known weights or relative precisions. Cases with large w_i are relatively more variable than are cases with small w_i ($i=1,2,\dots,n$).

We can transform the above model into a model in which the variances are equal, i.e.

$$Y^* = X^* \beta + \mathcal{E}^* \quad (1.10)$$

where $Y^* = W^{\frac{1}{2}} Y$, $X^* = W^{\frac{1}{2}} X$ and $\mathcal{E}^* = W^{\frac{1}{2}} \mathcal{E}$ such that

$$E(\mathcal{E}^*) = 0$$

$$Cor(\mathcal{E}^*) = \sigma^2 I$$

The least squares estimates in the transformed model can be obtained by minimizing the weighted sum of squares.

$$\begin{aligned} Q_w(\beta) &= (W^{\frac{1}{2}} Y - W^{\frac{1}{2}} X \beta)' (W^{\frac{1}{2}} Y - W^{\frac{1}{2}} X \beta) \\ &= Y' W Y - 2 \beta' X' W Y + \beta' (X' W X) \beta. \end{aligned} \quad (1.11)$$

Differentiating above w.r.t. β and equating to zero, we obtain

$$\hat{\beta}_w = (X'WX)^{-1}X'WY. \quad (1.12)$$

where β_w denotes the weighted least squares estimate of β , assuming again X to be of full rank.

Since these estimates minimize the weighted sum of squares,

$$Q_w(\beta) = \sum_i w_i (Y_i - \sum_j X_{ij} \beta_j)^2$$

β_w is called the weighted least squares estimate of β . In this sum of squares, the observations are weighted in proportion to the reciprocal of their variances. Since weighted least squares is nothing but least squares on the transformed observations, analytical properties of β_w are similar to those of $\hat{\beta}$. When the weights in W are unknown, the analytical distribution of $\hat{\beta}_w$ and $\hat{\sigma}_w^2$ are the same if the elements of W can be estimated consistently resulting in \hat{W} .

1.3.3 Method of Maximum Likelihood

Another general method of estimation is the method of Maximum Likelihood Estimation (m.l.e.). For this method we need to first define the likelihood function.

Definition 1.1

Let Y_1, Y_2, \dots, Y_n be observed random variables from a population with parameter θ . Then the likelihood function L is defined, in this case where the Y_i come

from an uncountable sample space with probability density function $f(Y; \theta)$, by

$$L(\theta; Y_1, Y_2, \dots, Y_n) = \prod_{i=1}^n f(Y_i; \theta)$$

and for countable sample spaces by

$$L(\theta; Y_1, Y_2, \dots, Y_n) = \prod_{i=1}^n P(Y_i; \theta)$$

where $P(Y_i; \theta)$ is the probability mass function. We shall note that the likelihood function is a function of θ given Y_1, Y_2, \dots, Y_n .

The principle of maximum likelihood consists of finding a value of θ which maximizes the likelihood function. Assuming that the likelihood function is a positive differentiable function of θ and the maximum does not occur on the boundary of the set of all admissible θ , we attempt to find a solution to the likelihood equation

$$\frac{\partial L}{\partial \theta} = 0 \quad (1.13).$$

Since $\log L$ is a monotonic function of L , L and $\log L$ attain their extreme values (maxima or minima) at the same value of θ . For the properties of m.l.e., one could see Rao [10].

1.4 THE INVERSE GAUSSIAN DISTRIBUTION AND ITS PROPERTIES

A random variable Y is distributed as inverse Gaussian with parameters μ and λ , i.e. $Y \sim \text{IG}(\mu, \lambda)$, if Y has the pdf given by

$$f(y; \mu, \lambda) = \left(\frac{\lambda}{2\pi y^3} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\lambda(y - \mu)^2}{(2\mu^2 y)} \right\} \quad y > 0, \quad (1.16)$$

where μ and λ are positive real values

The distribution was first derived by Schnodinger[11] as the first passage distribution of Brownian Motion with positive drift and absorbing barrier. It was later on derived as a limiting form of distribution of sample size in certain Sequential Probability Ratio Tests (SPRT) by Wald[14]. But, it was Tweedie[13] who first investigated its basic characteristics in detail.

The first two cumulants, namely the mean and the variance are given by

$$E(Y) = \mu, \text{Var}(Y) = \frac{\mu^3}{\lambda} \quad (1.17)$$

It can be seen that μ and λ are only partially interpretable as location and scale parameters. There is a remarkable relationship between positive and negative moments given by

$$E(Y^{-r}) = \frac{E(Y^{r+1})}{\mu^{2r+1}} \quad (1.18)$$

Schnodinger [11] showed that for a random sample Y_1, Y_2, \dots, Y_n where $Y_i \sim IG(\mu, \lambda)$, the maximum likelihood estimate of μ and λ are given by $\hat{\mu} = \bar{Y}$ and $\hat{\lambda} = n \sum_{i=1}^n (\frac{1}{Y_i} - \frac{1}{\bar{Y}})^{-1}$ where $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$.

Tweedie[13] proved that $Y \sim IG(\mu, n\lambda)$ and that $\lambda \sum (\frac{1}{Y_i} - \frac{1}{\bar{Y}}) \sim \chi_{n-1}^2$ distribution. Independence of Y and $\hat{\lambda}$ is fairly easily established by finding the conditional moment generating function of $\lambda \sum_{i=1}^n (\frac{1}{Y_i} - \frac{1}{\bar{Y}})$ given $\bar{Y} = y$. It is that of a χ^2 with $(n-1)$ degrees of freedom for all y . Therefore independence follows.

The statistics Y and $\sum_{i=1}^n (\frac{1}{Y_i} - \frac{1}{\bar{Y}})$ jointly form a complete sufficient statistic for (μ, λ)

The distribution of the reciprocal of the IG variable, i.e., the distribution of $Z = \frac{1}{Y}$ was also considered by Tweedie [13]. The p.d.f. of $Z = \frac{1}{Y}$ is

$$f(z) = \left(\frac{\lambda}{2\pi z}\right)^{\frac{1}{2}} \exp\left[\frac{-\lambda z}{2} + \frac{\lambda}{\mu} - \frac{\lambda}{2\mu^2 z}\right]$$

$$0 < z < \infty. \quad (1.19)$$

The mean and the variance of $\frac{1}{Y}$ which are easily obtained from (1.18) are given by

$$E\left(\frac{1}{Y}\right) = \frac{1}{\mu} + \frac{1}{\lambda}, \quad (1.20)$$

$$var\left(\frac{1}{Y}\right) = \frac{1}{\lambda\mu} + \frac{2}{\lambda^2}. \quad (1.21)$$

Chapter 2

RESAMPLING INFERENCE IN A STANDARD LINEAR MODEL

2.1 Introduction

Two methods have become very popular in recent years in investigating inference problems with increased accuracy as compared to the asymptotic methods, namely, the Jackknife Method and the Bootstrap Method. The Jackknife Method was originally proposed as a bias reducing technique by Quenouille[9] but has found itself a prominent place in statistical literature after the suggestion of Tukey[12]. It has been variously generalized and studied (see Gray and Schucany [6]).

This method has created a renewed interest in itself after the introduction of the Bootstrap Method proposed by Efron[4]. Because, they are viewed as competitors.

Our purpose in this thesis is to apply and compare these two methods in the case of inverse-Gaussian regression, which resembles somewhat to the general linear model given in section (1.2). Hence we first outline these methods in section (2.2) for inference on a single parameter and then review their generalization for the standard linear model in section (2.3).

2.2 Jackknife and Bootstrap For a Single Parameter

2.2.1 The Jackknife Method

We sample independent and identically distributed random quantities Y_1, Y_2, \dots, Y_n according to some unknown probability distribution F . Having observed $Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n$, we compute some statistic of interest, say

$$\theta = \theta(y_1, y_2, \dots, y_n). \quad (2.1)$$

Quenouille's [9] method is based on sequentially deleting point y_i , and recomputing $\hat{\theta}$. Removing point y_i from the data set gives a different empirical probability distribution, $\hat{F}(i)$ which assigns mass $\frac{M_i^*}{n-1}$ at each of the points $y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_n$, where M_i^* is the number of times Y_i appears in the bootstrap sample. The corresponding recomputed value of the statistic is

$$\hat{\theta}_{(i)} = \theta(\hat{F}_{(i)}) = \theta(y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_n) \quad (2.2)$$

Jackknife Estimate of Bias :

If

$$\theta_{(i)} = \frac{1}{n} \sum_{i=1}^n \hat{\theta}_{(i)}. \quad (2.3)$$

Quenouille's estimate of bias is

$$\widehat{BIAS} \equiv (n-1)(\hat{\theta}_{(i)} - \hat{\theta}). \quad (2.4)$$

leading to the bias corrected "jackknife estimate" of θ

$$\hat{\theta} - \widehat{BIAS} = n\hat{\theta} - (n-1)\hat{\theta}_{(.)}. \quad (2.5)$$

The usual rationale for \widehat{BIAS} and $\tilde{\theta}$ goes as follows. If E_n denotes the expectation for sample size n , $E_n \equiv E_F \hat{\theta}(Y_1, Y_2, \dots, Y_n)$, then for many common statistics, including most *maximum likelihood estimates*,

$$E_n = \theta + \frac{a_1(F)}{n} + \frac{a_2(F)}{n^2} + \dots, \quad (2.6)$$

where the functions $a_1(F), a_2(F), \dots$ do not depend upon n ; see Hinkley [7]. Note that

$$E_F \hat{\theta}_{(.)} = E_{n-1} = \theta + \frac{a_1(F)}{n-1} + \frac{a_2(F)}{(n-1)^2} + \dots, \quad (2.7)$$

and therefore

$$E_F \tilde{\theta} = nE_n - (n-1)E_{n-1} = \theta - \frac{a_2(F)}{n(n-1)} + a_3(F)\left(\frac{1}{n^2} + \frac{1}{(n-1)^2}\right) + \dots. \quad (2.8)$$

We see that $\tilde{\theta}$ has a bias of order $O(\frac{1}{n^2})$ compared to $O(\frac{1}{n})$ for the original estimator.

Tukey [12] further suggested that using the quantities $P_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)}$ as i.i.d. with expectation θ , which gave popularity to the jackknife method.

2.2.2 The Bootstrap Method

Efron[4] advocates the use of a sampling with replacement method in statistical computations, and this gave birth to the concept known as the "bootstrap". It was presented as a refinement of Quenouille - Tukey jackknife. In bootstrapping, only

the observed data are required and no other extraneous data are needed. Thus, one is in fact pulling oneself up by the bootstraps. This is how the "bootstrap" gets its name.

The bootstrap is conceptually the simplest of all the techniques. This method provides a useful tool for estimating the sampling properties of a given statistic. This is done without prior knowledge of the parent distribution of an observed sample. To a certain extent, it is for this reason that both the jackknife and the bootstrap methods are often referred to as "distribution - free" methods. A statistical method is distribution-free provided its application is valid, regardless of the underlying distribution. However, the accuracy of Efron's bootstrap does vary with the class of statistics and with the underlying probability distribution.

Given a statistic $\theta(Y_1, Y_2, \dots, Y_n)$ defined symmetrically in Y_1, Y_2, \dots, Y_n i.i.d. $\sim F$, we write the standard deviation of θ as

$$sd = \sigma(F, n, \theta) = \sigma(F). \quad (2.9)$$

Thus the last notation emphasizes the fact that, given the sample size n and the form of the statistic $\theta(\dots)$, the standard deviation is a function of the unknown probability distribution F .

The bootstrap estimate of the standard deviation is simply $\sigma(\cdot)$ evaluated at $F = \hat{F}$,

$$\widehat{SD} = \sigma(\hat{F}) \quad (2.10)$$

Since \hat{F} is the non-parametric maximum likelihood estimate of F , another way to express(2.10) is that \widehat{SD} is the non- parametric m.l.e. of sd (if σ is a one to one

function of F).

Monte Carlo Evaluation of \widehat{SD} :

Usually the function $\sigma(F)$ cannot be written down explicitly. In order to carry out the calculation of \widehat{SD} in (2.10), it is then necessary to use a Monte Carlo algorithm.

(i) Fit the nonparametric m.l.e. of F .

$$F(y) = \frac{\text{number of observations} \leq y}{n} \quad (2.11)$$

(ii) Draw a "bootstrap sample" from F .

$$Y_1^*, Y_2^*, \dots, Y_n^* \text{ i.i.d. } \sim \hat{F}, \quad (2.12)$$

and calculate $\theta^* = \theta(Y_1^*, Y_2^*, \dots, Y_n^*)$.

(iii) Repeat step (ii) a large number of times B , obtaining "bootstrap replications" $\theta^{*1}, \theta^{*2}, \dots, \theta^{*B}$, and calculate

$$\widehat{SD} = \left\{ \frac{1}{B-1} \sum_{b=1}^B (\theta^{*b} - \hat{\theta}^*)^2 \right\}^{\frac{1}{2}} \quad (2.13)$$

where $\hat{\theta}^* = \sum_{b=1}^B \theta^{*b} / B$.

As $B \rightarrow \infty$, \widehat{SD} converges a.s. to SD in (2.10). In practice B is chosen to be large.

Similar algorithms produce estimators of the Bias and MSE of $\hat{\theta}$.

2.3 Jackknife and Bootstrap For Regression parameters in a Linear Model

2.3.1 Jackknife Method

Consider a linear regression model given by

$$Y = X\beta + \mathcal{E}, \quad (2.14)$$

where, Y is an $(n \times 1)$ observable random vector; X is an $(n \times p)$ matrix (such that $\lim_{n \rightarrow \infty} (\frac{X'X}{n})$ is finite), β is a $(p \times 1)$ vector of unknown regression coefficients and \mathcal{E} is an $(n \times 1)$ vector of disturbances which follow the usual assumptions, namely, $E(\mathcal{E}) = 0$ and $E(\mathcal{E}'\mathcal{E}) = \sigma^2 I$.

The statistic of interest is the estimator β .

$$\beta = (X'X)^{-1}X'Y \quad (2.15)$$

We assume that the $n \times p$ matrix X is of full rank, so that the matrix $(X'X)$ has an inverse.

The basic components of the standard jackknife procedure are parameter estimates obtained by successively deleting single observation. Following Quenouille [9], Tukey[12], the corresponding estimator of β obtained by deleting the i th observation from the sample data is,

$$\hat{\beta}_{-i} = (X'_{-i}X_{-i})^{-1}X'_{-i}Y_{-i} \quad (2.16)$$

where X_{-i} is an $(n-1) \times p$ matrix of X variables, and Y_{-i} is the $(n-1) \times 1$ vector (Y less the i th element). Following Hinkley [7], (2.16) can be written as

$$\beta_{-i} = \beta - \frac{(X'X)^{-1}X_i\hat{\mathcal{E}}_i}{1-w_i} \quad (2.17)$$

where $\hat{\mathcal{E}}_i$ is the residual $Y_i - X_i'\beta$ and $w_i = X_i'(X'X)^{-1}X_i$ is the i -th element of the projection matrix $X(X'X)^{-1}X'$ onto the column space of X , i.e. the $p \times 1$ vector of the i th observation on p regressors.

To describe the standard Jackknife procedure, we define

$$P_i = n\beta - (n-1)\beta_{-i} \quad (i = 1, 2, \dots, n) \quad (2.18)$$

for which the Jackknife estimator is given by

$$\hat{\beta} = \frac{1}{n} \sum_i^n P_i. \quad (2.19)$$

Using (2.17) and (2.18) in (2.19), we obtain

$$\hat{\beta} = \beta + (n-1)(X'X)^{-1} \sum_i^n (1-w_i)^{-1} X_i \hat{\mathcal{E}}_i. \quad (2.20)$$

Quite generally the jackknife estimator removes the bias of order (n^{-1}) . Here $\hat{\beta}$ is unbiased, and therefore this property is redundant. Clearly $\hat{\beta}$ is unbiased, since $E(\hat{\mathcal{E}}_i) = 0$, so the fact that $\hat{\beta}$ and β are generally different, together with the Gauss-Markov property implies that

$$\text{var}(\hat{\beta}) > \text{var}(\beta)$$

The exceptions to this occur in balanced linear models, where w_i is constant. A somewhat weaker property of β is general consistency, which holds if $n^{-1}(X'X)$ converges to a positive definite matrix, this implying $\max w_i \rightarrow 0$; see Miller[8].

Similarly we can obtain $\sigma_{\beta_i}^2$, a corresponding estimator of σ^2 .

However, if (2.14) is an unbalanced model with the lack of balance reflecting through the "distances" w_i , Hinkley[7] therefore proposed the weighted jackknife procedure for which the pseudo values are defined as,

$$Q_i = \beta + n(1 - w_i)(\beta - \beta_{(i)}) \quad (i = 1, 2, \dots, n). \quad (2.21)$$

The corresponding weighted jackknife estimator is given by

$$\beta_w = \frac{1}{n} \sum_i^n Q_i = \beta \quad (2.22)$$

and the estimator of variance covariance matrix of β is given by

$$\begin{aligned} V_w &= \{n(n-p)\}^{-1} \sum_i^n (Q_i - \beta_w)(Q_i - \beta_w)' \\ &= n(n-p)^{-1} (X'X)^{-1} \left(\sum \mathcal{E}^2 X_i X_i' \right) (X'X)^{-1}; \end{aligned} \quad (2.23)$$

where in each case the explicit form for the linear model is given. The denominator $(n-p)$ in V_w reflects the number of degrees of freedom in the residual vector, and makes V_w exactly unbiased in the balanced case when $w_i = pm^{-1}$.

2.3.2 The Bootstrap Method :

The assumptions underlying the model (2.14), as described in section (2.3.1) of Freedman [5] are :

A.2.1 : For the $n \times p$ matrix X ,

$$\lim \frac{X'X}{n} \rightarrow M(finite), \text{ a.e.} \quad (2.24)$$

the $p \times p$ matrix M being positive definite (p.d.) of finite elements. This assumption excludes variables which grow indefinitely, as n goes to infinity (see White [15])

A.2.2 . The components of \mathcal{E} are i.i.d. with common unknown distribution function, F , having mean zero and constant variance.

Substituting (2.14) in (2.15), we get

$$\hat{\beta} = \beta + (X'X)^{-1}X'\mathcal{E} \quad (2.25)$$

The predictor of \mathcal{E} is $\hat{\mathcal{E}}$, given by

$$\hat{\mathcal{E}} = Y - X\hat{\beta} = M\mathcal{E} \quad (2.26)$$

A bootstrap algorithm for the linear regression model, whose purpose is to generate $\beta_{(j)}^*$ for $j=1,2,\dots,B$, is described in what follows:

Algorithm 2.1

Purpose : To obtain $\beta_{(j)}^*$ for $j=1,2,\dots,B$ where B is large, usually between 200 and 10,000.

Step 1 : Obtain $\hat{\beta}$ and \mathcal{E} according to the specification of (2.25) and (2.26) respectively.

Step 2 : Generate n random integers, drawn with replacement from the set $\{1,2, \dots,n\}$.

Step 3 : Reconstruct the linear responses Y^* 's as

$$Y_{(j)}^* = X\hat{\beta} + \hat{\mathcal{E}}_{(j)} \quad (2.27)$$

Step 4 : Compute

$$\beta_{(j)}^* = (X'X)^{-1}X'Y_{(j)}^*. \quad (2.28)$$

Step 5 : Repeat steps 2,3 and 4 for $j=1,2,\dots,B$.

Asymptotic theorem by Freedman [5] :

As a consequence of A.2.1 and A.2.2, $\hat{\beta}$ is a consistent estimate of β and

$$\sqrt{n}(\hat{\beta} - \beta) \quad (2.29)$$

is asymptotically normal having mean zero and variance $\sigma^2 M_{XX}^{-1}$ (see White[15], P14). Also, an unbiased estimator of σ^2 is given by

$$\sigma_n^2 = (n - k)^{-1} \mathcal{E}' \mathcal{E} \quad (2.30)$$

and $n^{-1} \mathcal{E}' \mathcal{E}$ is a maximum likelihood estimator of σ^2 , provided that the errors are normally distributed.

The Bootstrap Estimate Of Bias :

Let

$$E_F[\mathcal{H}(F)] = \mathcal{H}(F) + Bias$$

i.e.,

$$Bias = E_F[\mathcal{H}(F)] - \mathcal{H}(F) \quad (2.31)$$

and

$$V[\mathcal{H}(F)] = MSE[\mathcal{H}(F)] - (Bias)^2. \quad (2.32)$$

Suppose that we wish to estimate the bias of a functional statistic given by (2.31).

We can take $R(Y, F) = \mathcal{H}(\hat{F}) - \mathcal{H}(F)$, and use the bootstrap algorithm to estimate

$E_F R = Bias$. In this case

$$R^* = R(Y^*, \hat{F}) = \mathcal{H}(\hat{F}^*) - \mathcal{H}(\hat{F}) = \hat{\beta}^* - \hat{\beta} \quad (2.33)$$

where $\beta^* = \beta(F^*)$, F^* being the empirical probability distribution of the bootstrap sample: F^* puts mass $\frac{M_i^*}{n}$ on Y_i , where M_i^* is the number of times Y_i appears in the bootstrap sample.

If the bootstrap distribution for the estimator $\hat{\beta}$ is biased, then the bias can easily be corrected by the following two methods :

(i) Nominal Bias Correction : In this method bias correction can be done by the formula

$$\hat{\beta} = \beta - (\beta^* - \hat{\beta}) = 2\hat{\beta} - \hat{\beta}^* \quad (2.34)$$

(ii) Percentage Bias Correction : Let us define

$$\gamma = \frac{\hat{\beta} - \beta}{\hat{\beta}}$$

such that

$$\gamma = \frac{\hat{\beta}^* - \hat{\beta}}{\hat{\beta}}$$

and

$$1 + \gamma = \frac{\hat{\beta}^*}{\hat{\beta}}$$

which implies that

$$\begin{aligned} \beta &= (1 + \gamma)^{-1} \hat{\beta} \\ &= \frac{\hat{\beta}^2}{\hat{\beta}^*} \end{aligned} \quad (2.35)$$

In the literature, difficulties associated with the ordinary least squares residuals for bootstrapping the Linear Model have been partly attributed to the fact that, although $\hat{\mathcal{E}}$ is asymptotically an appropriate estimate of \mathcal{E} , it has a covariance matrix which depends on the design matrix X .

In the IG reciprocal model, the Least Squares Estimators are not the same as the Maximum Likelihood Estimators and whereas the (finite sample) sampling properties of the least squares estimators are easier to obtain than those of the m.l.e., the former are less efficient for the large samples.

In this thesis we are using m.l.e. to estimate the regression coefficients. Moreover, the maximum likelihood estimates corresponds to Weighted Least Squares estimators with response variable $1/Y$ and associated weight Y (Whitmore [16]) which makes the adaptation of the bootstrap method possible for the IG regression model.

Chapter 3

RESAMPLING INFERENCE IN AN INVERSE GAUSSIAN REGRESSION MODEL

3.1 Introduction

In this chapter we develop the resampling procedure for the inverse Gaussian model. The pdf of the IG distribution is given by the equation (1.16) as discussed in the Chapter 1.

The deviation of the IG distribution can be cast in the context of fatigue growth or accumulation of fatigue or damage over a period of time according to a Weiner process.

In this thesis we consider a reciprocal linear model for the mean of the inverse Gaussian distribution, as proposed by Bhattacharya and Fries [3], in the context of accumulation of fatigue. They consider the following situation : assume that n similar items are subject to stress at different levels x_1, x_2, \dots, x_n until they break. Each item has a characteristic breaking threshold, which we assume to be the same for all items. The different stress levels, however, imply that the accumulation of

fatigue proceeds for different items at different levels. We may then model the drift as a function of the covariate x , if the accumulation follows a Brownian motion ; and when this drift is linear in x , the time of breaking for the i th item follows an inverse-Gaussian distribution with mean μ_i where

$$\frac{1}{\mu_i} = \beta_0 + \beta_1 x_i$$

and, because the breaking threshold is the same for all items, we assume a constant λ .

In this chapter, we develop Jackknife and Bootstrap methods, as discussed in Chapter 2, for estimating the parameter β and its bias, in section 3.2 and 3.3 respectively.

3.1.1 The Model

We assume that at each design point $X_i \in \mathbb{R}^p, i = 1, 2, \dots, k$ there are independent observations $Y_{ij}, j = 1, 2, \dots, n_i$. Suppose that $Y_{ij} \sim IG(\mu_i, \lambda)$. We reparametrize for convenience (μ, λ) to (μ, σ) where $\sigma = \frac{1}{\lambda}$. The assumption of a linear drift of an inversely proportional breaking time is equivalent to assume a linear model for the reciprocal of the mean i.e.

$$\mu_i^{-1} = X_i' \beta \quad , i = 1, 2, \dots, k \quad (3.1)$$

where $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ is a vector of regression parameters and $X_i = (x_{i1}, \dots, x_{ip})$ is a vector of explanatory variables. Note that $x_{i1} = 1$ for a non zero intercept model.

Denoting the $N \times p$ matrix of observations of the explanatory variables by X , we get the following linear model for $\delta = (\mu_1^{-1}, \mu_2^{-1}, \dots, \mu_k^{-1})$:

$$\delta = X\beta \quad (3.2)$$

where $N = \sum_{i=1}^k n_i$.

Assume also that $k \leq p + 1$. We now introduce the following notations :

$$Y_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}. \quad (3.3)$$

$$Y = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}}{N}, \quad (3.4)$$

$$R = \frac{\sum_i \sum_j^{n_i} Y_{ij}^{-1}}{N}, \quad (3.5)$$

$$V = \sum_i \sum_j (Y_{ij}^{-1} - Y_i^{-1}), \quad (3.6)$$

$$X = \frac{\sum_{i=1}^k n_i X_i}{N}, \quad (3.7)$$

$$C = \text{diag}(n_1, n_2, \dots, n_k), \quad (3.8)$$

$$D = \text{diag}(Y_1, Y_2, \dots, Y_k), \quad (3.9)$$

$$X' = (X_1, X_2, \dots, X_k), \quad (3.10)$$

and

$$S = X' C D X. \quad (3.11)$$

Here D and S are matrix statistics.

Some important properties of the inverse-Gaussian which will be used here are

(i) Y_i and $\sum_{j=1}^{n_i} (Y_{ij}^{-1} - Y_i^{-1})$ are independent.

(ii) $Y_i \sim IG(\mu_i, \frac{1}{n\lambda})$, and

(iii) $\lambda \sum_{j=1}^{n_i} (Y_{ij}^{-1} - Y_i^{-1}) \sim \chi^2_{(n-1)}$.

Moreover

$$E(Y_i^{-1}) = \mu_i^{-1} + \frac{1}{n\lambda} \quad (3.12)$$

$$Var(Y_{ij}^{-1}) = \mu_i^{-1} \frac{\sigma}{n} + 2\left(\frac{\sigma}{n}\right)^2. \quad (3.13)$$

The proof of these results can also be found in [3].

3.1.2 Maximum Likelihood Estimation

The likelihood function is given here by

$$L(\mu, \sigma) = \prod_{i=1}^k \prod_{j=1}^{n_i} f(Y_{ij}; \mu_i, \sigma) \quad \sigma > 0 \quad (3.14)$$

and is propotional to

$$\begin{aligned} \sigma^{-\frac{n}{2}} \prod_{i=1}^k \prod_{j=1}^{n_i} \exp\left\{-\frac{Y_{ij}^{-1}(Y_{ij}\mu_i^{-1}-1)^2}{(2\sigma)}\right\} \\ = \sigma^{-\frac{n}{2}} \exp\left\{-\frac{Q(\beta)}{(2\sigma)}\right\} \end{aligned} \quad (3.15)$$

where

$$\begin{aligned} Q(\beta) &= \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^{-1} (Y_{ij} X_{ij}' \beta - 1)^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} \{ (Y_{ij} (X_{ij}' \beta))^2 - 2X_{ij}' \beta \} + \sum_{i=1}^k \sum_{j=1}^{n_i} Y_{ij}^{-1} \end{aligned} \quad (3.16)$$

The above may further be simplified as

$$Q(\beta) = \sum_{i=1}^k \{ n_i Y_i (X_i' \beta)^2 - 2n_i X_i' \beta \} + \sum_{i=1}^k \sum_j Y_{ij}^{-1}$$

$$\begin{aligned}
&= \sum_{i=1}^k n_i \{Y_i(X'_i \beta)^2 - 2X'_i \beta\} + \sum_i \sum_j Y_{ij}^{-1} \\
&= (DX\beta - 1)'C'D^{-1}(DX\beta - 1) + V
\end{aligned} \tag{3.17}$$

where V and D are defined in (3.6) and (3.9) respectively.

Now referring to (3.14) and (3.17), the partial derivatives of log L with respect to β and σ yield a set of likelihood equations, namely,

$$N\sigma - Q(\beta) = 0 \tag{3.18}$$

and

$$S\beta - V'C1 = 0 \tag{3.19}$$

The solution of (3.18) and (3.19) is given by

$$\beta = NS^{-1}X = (X'1X)^{-1}X'1 \tag{3.20}$$

$$\sigma = R - \hat{\beta}'X \tag{3.21}$$

Even though the estimators $\hat{\beta}$ and $\hat{\sigma}$ are obtained as unique solutions of the likelihood equations, they are not necessarily the maximum likelihood estimators because the estimated mean $(X'\hat{\beta})'$ at a given design point X may turn out to be negative. Therefore we can call these estimators of β and σ as Pseudo m.l.e of β and σ , respectively. However, unless N is very small, this is not a serious problem, because the theorem proved by Bhattacharya and Fries [3] given below shows that $\hat{\beta}$ and $\hat{\sigma}$ serve the primary goal of m.l.e's which is to provide an efficient likelihood estimator.

Theorem: As $N \rightarrow \infty$, $n_i/N_i \rightarrow h_i > 0$ and $N^{-1}C \rightarrow H$, we have

$$a) \sqrt{N}(\hat{\beta} - \beta) \rightarrow N_p(0, \sigma \Delta^{-1}),$$

b) $\sqrt{N}(\frac{\hat{\sigma}}{\sigma} - 1) \rightarrow N(0, 2)$.

c) $\hat{\beta}$ and $\hat{\sigma}$ are asymptotically independent.

where $\Delta = X' H M X$, $M = \text{diag}(\mu_1, \mu_2, \dots, \mu_k)$ and $H = \text{diag}(h_1, h_2, \dots, h_k)$.

Bias of the Estimator β : In finite samples $\hat{\beta}$ may have a "large" bias. To evaluate this bias, we use a two term matrix inverse expansion to get

$$E(\hat{\beta}) = E((X'YX)^{-1}X'1) = \beta + F^{-1}X'R\Delta R X F^{-1}X'1 \quad (3.22)$$

where $\Delta = \text{diag}(c_n \frac{\mu_i}{\lambda})$, $C = \lambda F^{-1}X'R = Y - M$ and $F = X'MX$ (for detailed derivation see the appendix).

The bias is thus of order $O(\frac{1}{n})$ but depends upon β and λ . Thus it is of interest to estimate the finite sample distribution of $\hat{\beta}$ and/or to eliminate bias of $\hat{\beta}$.

The following illustrated techniques are used to estimate bias of $\hat{\beta}$ when the model observations are coming from inverse Gaussian distribution.

3.2 Jackknife Method

From section 3.1 it is clear that $\hat{\beta}$ has a bias of order $O(1/n)$. Now we use Jackknife method to estimate this bias as discussed in chapter 2.

In our case for the model as described in section (1.2), analogous to (2.15),(2.17),(2.18) and (2.20) :

$$\beta = (X'YX)^{-1}X'1. \quad (3.23)$$

$$\hat{\beta}_{-i} = \hat{\beta} - \frac{(X'YX)^{-1}X_i(1 - Y_i\hat{Y}_i)}{(1 - Y_iw_i)}, \quad i = 1, 2, \dots, p \quad (3.24)$$

$$P_i = \hat{\beta} + (n-1)(X'YX)^{-1}X_i(1 - Y_i\hat{Y}_i)(1 - Y_iw_i)^{-1} \quad (3.25)$$

and

$$\hat{\beta} = \hat{\beta} + \frac{(n-1)}{n} \sum_{i=1}^n (1 - Y_iw_i)^{-1}(1 - Y_i\hat{Y}_i)X_i \quad (3.26)$$

where $w_i = X_i(X'YX)^{-1}X_i'$ and $Y_i = X_i'\beta$.

Taking expectation, we get

$$E(\hat{\beta}) = \beta + O\left(\frac{1}{n}\right) + \frac{(n-1)}{n}(X'YX)^{-1} \sum_{i=1}^n E\{(1 - Y_iw_i)^{-1}(1 - Y_i\hat{Y}_i)\}X_i \quad (3.27)$$

The summation needs tedious computations but we notice that it may depend upon

β and λ and that it is difficult to assess the order of the bias in this particular case.

Analogous to weighted jackknife estimates given by the equations (2.21) and (2.22),

we obtain here

$$Q_i = \hat{\beta} + n(\hat{\beta} - \hat{\beta}_{-i})(1 - Y_iw_i), \quad i = 1, 2, \dots, p \quad (3.28)$$

and

$$\hat{\beta}_w = n^{-1} \sum_{i=1}^n Q_i = \hat{\beta} + \sum_{i=1}^n (1 - Y_iw_i)(\hat{\beta} - \hat{\beta}_{-i}) \quad (3.29)$$

where $w_i = X_i'(X'YX)^{-1}X_i$ and $Y_i = X_i'\beta$.

Hence, substituting (3.24) in (3.29), we get

$$\hat{\beta}_w = \hat{\beta}$$

and taking expectations of terms of order $\frac{1}{n}$ we have

$$E(\hat{\beta}_w) = \beta + O\left(\frac{1}{n}\right), \quad (3.30)$$

since from equation (3.22) it is clear that the bias of $\hat{\beta}$ is of order $O(\frac{1}{n})$ but depends upon β and λ .

In this particular case, we can say that the jackknife procedure fails to remove the bias of $\hat{\beta}$. In the next section we use the "Bootstrap" procedure to estimate the bias of the parameter β .

3.3 Bootstrap Method

The bootstrap method is similar to the Monte Carlo Method, except that in the bootstrap experiments, $\mathcal{E}_{(j)}^*$ is drawn from \hat{F} instead of F . The reason for using \hat{F} is that it is attainable, whereas F is neither known nor observable.

Eventhough the m.l.e. $\hat{\beta}$ is the weighted least squares estimate, adoption of the standard bootstrap method is not directly applicable. We consider the following two methods to adopt the bootstrapping procedure for the IG regression model;

Method 1: Let the transformed residuals be given by

$$\mathcal{E}_i = (X_i' \hat{\beta})^{-2} \mathcal{E}_i \quad , i = 1, 2, \dots, n \quad (3.31)$$

where the $\hat{\mathcal{E}}_i$'s are drawn from \hat{F} . The bootstrap algorithm for the linear model using rescaled residuals (3.31) can be described as follows :

Algorithm 3.1

Purpose : The purpose of this algorithm is the same as that of algorithm 2.1, except that \mathcal{E} is now replaced by \mathcal{E} whereas step 3 is replaced by

Step 3 : Reconstruct the linear responses as

$$Y_{(j)}^* = R_{(j)}^* + \tilde{\mathcal{E}}_{(j)} \quad (3.32)$$

where $R_{(j)l}^* = \frac{1}{N_l^*} \quad l = 1, 2, \dots, n$.

Repeat steps 2,3 and 4 for $j=1,2,\dots,B$.

Method 2: Let us define the transformed residuals as

$$\mathcal{E}_l = \sqrt{N_l^*} \tilde{\mathcal{E}}_l; \quad (3.33)$$

The bootstrap algorithm to generate $\hat{\beta}_{(j)}^*$ in this case is as follows :

Algorithm 3.2

Purpose : The purpose of this algorithm is the same as that of algorithm

2.1 except that \mathcal{E} is now replaced by $\tilde{\mathcal{E}}$ whereas step 3 is replaced by

Step 3 : Reconstruct the linear responses as

$$R_{(j)}^* = X\beta + \tilde{\mathcal{E}}_{(j)} \quad (3.34)$$

where $R_{(j)l}^* = \frac{1}{N_{(j)l}^*} \quad j = 1, 2, \dots, B; \quad l = 1, 2, \dots, n$ Note that the sample mean of $\tilde{\mathcal{E}}$ (either obtained by method 1 or method 2) is zero.

Based on results for the bootstrap standard linear model we expect that :

For large n ,

$$\sqrt{n}(\hat{\beta} - \beta) \rightarrow \sqrt{n}(\hat{\beta}^* - \hat{\beta}) \quad (3.35)$$

where $\hat{\beta}_{(j)}^*$ is bootstrap estimator obtained by method 1 or method 2. A rigorous proof of (3.35) was given by Babu [1].

Estimate of Bias : As explained in chapter 2, the bootstrap estimate of the bias is approximated by

$$BIAS = \frac{1}{B} \sum_{b=1}^B \beta_{(b)}^* - \hat{\beta} = \hat{\beta}^* - \hat{\beta} \quad (3.36)$$

Estimate of the MSE of $\hat{\beta}$

Because of (3.35) the MSE of $\hat{\beta}$ for large samples may be estimated by

$$\widehat{MSE}(\hat{\beta}) = \frac{1}{B} \sum_{b=1}^B (\beta_{(b)}^* - \hat{\beta})(\beta_{(b)}^* - \hat{\beta})' \quad (3.37)$$

a numerical study of these is given in the next chapter.

Chapter 4

A NUMERICAL STUDY OF VARIOUS ESTIMATORS

4.1 Design of the Simulation Study

In this chapter we will apply our methods to estimate the parameters in the inverse Gaussian linear regression model. In a linear model the true disturbances are unobservable with empirical applications to real world data, but ordinary least squares (OLS) residuals can easily be computed. Freedman and Peters [5] advocate the use of either best linear unbiased residuals or inflated OLS residuals.

Here, we will use maximum likelihood estimates of the regression coefficient. Even though the m.l.e. of β is a weighted least squares estimate but adaptation of the standard bootstrap method in the case of general linear models is not directly applicable. We consider two methods to adapt the bootstrap method for IG regression model, as mentioned in Chapter 3 (section 3.3). The bootstrap residuals are computed according to these two methods and models are reconstructed by using algorithms 3.1 and 3.2 respectively.

Since m.l.e. β itself is biased in most of the cases, the regular bootstrap estimator is biased also and hence, is corrected by nominal correction and percentage correction, as discussed in Chapter 3. A simulation study is done to compare these

estimates based upon the regression models as described in Chapter 2. For the simulation design we consider $n_i = 1, \forall i = 1, 2, \dots, k$. When $p=2$, the intercept and slope of the regression line are assigned different sets of values. Both the disturbances and values for the independent variables are generated by IMSL RNUNF() subroutines which have been well tested and are very reliable. Note that the random number generator used in this algorithm should be chosen to have a very large period. The disturbances are calculated by method 1 and method 2 (as in section 3.3).

Given $\hat{\beta}$ and $\hat{\varepsilon}$ of (2.25) and (2.26), respectively, the bootstrap residuals and reconstructed responses are obtained. The observations of the independent variables are assumed to come from a Uniform (1,2) distribution. For each n , these observations are drawn once and then fixed throughout the experiment.

Separate experiments are constructed (for $n=10,20,30,40,50$), to determine the effects of the sample size on the different values of μ and λ . For each experiment, $NS=1000$ samples are considered. In each sample, $B = 500$ trials are constructed and three different sets of values of intercept, slope (i.e. $\beta_0 = 0, \beta_1 = 1.0; \beta_0 = 1.0, \beta_1 = 1.0$ and $\beta_0 = 1.0, \beta_1 = 0$), values of λ (5,10,20), and of μ (as it depends on β and X_i) are taken. In each trial, the variations (bias and MSE) of the regression coefficients are computed. At each trial, the bias and MSE of the regular bootstrap estimates are calculated. Average of bias estimates based on 1000 simulation trials are calculated. Both the bias and MSE are used as yardsticks for comparison among the four estimators (maximum likelihood, regular bootstrap, nominal bias corrected

bootstrap and percentage bias corrected bootstrap).

Detailed tables of the biases and MSE of the estimators based on the simulation are given in the Tables 4.1 to 4.3. Our conclusions about the bias and the MSE's of different estimators are given in the next section. This section further provides an assesment of the estimator of the MSE of m.l.e.'s.

4.2 Conclusions

4.2.1 The Bias of the Estimators

The results in Table 4.1 indicate that for small values of n and λ , the percentage bootstrap estimator (PBC) has a large bias, but as n increases and λ increases, the biases decrease and the nominal bootstrap correction estimator (NBC) and PBC tend to have smaller bias than that of m.l.e. Evenif with these increased values of n and λ , the bias in the bootstrap estimator is falling more repidly than that of the m.l.e., it stil comes out to be the best esimator in the sense of having the smallest bias.

The results in Table 4.2 indicate that for all values of n , as λ increases, the biases decrease. Specially for $n=20$ and $\lambda = 10$, the bias of the m.l.e. is negligible but the bootstrap estimators still have non-negligible bias. For $n = 50$, the regular bootstrap as well as the corrected bootstrap estimators have non-negligible bias.

Table 4.1 a

Biases of the Estimators Using Method 1

$$\beta_0 = 1, \quad \beta_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	-.0475	.1185	.0235	.2842
		β_1	.0641	.1323	-.0042	.0152
	10	β_0	-.0306	.0651	.0040	.2064
		β_1	.0379	.0775	-.0017	.0080
	20	β_0	-.0216	.0306	-.0125	-.0125
		β_1	.0235	.0379	.0091	.0120
20	5	β_0	.0130	.0505	.0246	.0288
		β_1	.0027	.0060	.0113	.0003
	10	β_0	.0008	.0211	.0196	.0162
		β_1	.0053	.0005	.0100	.0067
	20	β_0	-.0008	.0096	.0112	.0109
		β_1	.0035	.0013	.0058	.0056
30	5	β_0	.0124	.0358	-.0110	.0019
		β_1	-.0003	.0041	.0034	.0430
	10	β_0	.0029	.0149	-.0091	.0078
		β_1	.0022	.0003	.0040	.0142
	20	β_0	.0047	.0113	-.0019	.0016
		β_1	-.0008	.0017	.0000	.0008

Table 4.1 a (continued)

Biases of the Estimators Using method 1

$$j_0 = 1, \quad j_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	j_0	.0070	.0174	-.0034	-.0015
		j_1	.0021	.0041	.0000	.0041
	10	j_0	.0035	.0093	-.0024	-.0019
		j_1	.0018	.0027	.0009	.0011
	20	j_0	.0026	.0063	-.0012	-.0011
		j_1	.0001	.0005	-.0004	-.0001
50	5	j_0	.0210	.0324	.0096	.1394
		j_1	.0091	-.0092	-.0089	-.0093
	10	j_0	.0139	.0201	.0077	.0080
		j_1	-.0067	-.0067	-.0067	-.0067
	20	j_0	.0092	.0132	.0052	.0053
		j_1	-.0044	-.0044	-.0044	-.0049

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.1 b

Biases of the Estimators Using Method 1

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C'	P B C'
10	5	β_0	-.0301	.3353	.2751	.0217
		β_1	.0554	.2971	-.1864	.1264
	10	β_0	-.0368	-.2152	.1415	.3468
		β_1	.0436	.1911	.1039	-.0790
	20	β_0	-.0109	.0859	.0641	.0891
		β_1	.0166	.0820	.0488	.0428
20	5	β_0	.0072	-.0718	.0863	.0686
		β_1	.0071	.0798	-.0656	.0577
	10	β_0	.0000	-.0331	.0331	.0374
		β_1	.0063	.0396	-.0270	.0248
	20	β_0	-.0001	.0136	.0134	.0128
		β_1	.0027	.0188	-.0133	-.0127
30	5	β_0	.0037	.0529	.0602	.0371
		β_1	.0055	.0590	-.0481	.0437
	10	β_0	.0096	.0132	.0324	.0238
		β_1	-.0016	.0227	.0259	.0248
	20	β_0	.0038	.0058	.0133	.0137
		β_1	-.0002	.0123	-.0127	.0123

Table 4.1 b (continued)

Biases of the Estimators Using method 1

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.0069	-.0425	.0563	.0264
		β_1	.0031	.0493	-.0431	-.0400
	10	β_0	.0051	-.0170	.0272	.0311
		β_1	.0001	.0228	-.0229	-.0220
	20	β_0	.0021	-.0074	.0121	.0129
		β_1	-.0004	.0117	-.0126	-.0123
50	5	β_0	.0332	-.0035	.0700	.0637
		β_1	-.0167	.0198	-.0531	-.0514
	10	β_0	.0192	.0028	.0356	.0370
		β_1	-.0092	.0092	-.0277	-.0270
	20	β_0	.0167	.0100	.0235	.0238
		β_1	-.0099	.0000	-.0198	-.0196

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.1 c

Biases of the Estimators Using method 1

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	-.0185	.2350	.1980	.0347
		β_1	.0466	.1913	-.0981	.0279
	10	β_0	-.0178	.2561	.2209	.0093
		β_1	.0297	.2062	-.1468	.1033
	20	β_0	-.0134	.1932	.1664	.1020
		β_1	.0183	.1584	.1219	-.0972
20	5	β_0	.0162	-.1371	.1695	.0881
		β_1	.0013	.1199	-.1173	.0942
	10	β_0	.0081	-.0911	.1073	.0738
		β_1	.0011	.0813	-.0792	.0698
	20	β_0	-.0004	-.0490	.0182	.0384
		β_1	.0034	.0144	.0375	.0345
30	5	β_0	.0173	.1441	.1786	.0997
		β_1	-.0030	.1269	.1328	-.1128
	10	β_0	.0047	.0742	.0836	.0571
		β_1	.0006	.0662	-.0650	-.0587
	20	β_0	.0067	.0281	.0414	.0306
		β_1	-.0021	.0282	-.0323	.0307

Table 4.1 c (continued)

Biases of the Estimators Using method 1

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.0099	.1192	.1390	.0765
		β_1	.0014	.1040	-.1012	-.0890
	10	β_0	.0065	-.0617	.0746	.1027
		β_1	-.0004	.0561	-.0568	-.0524
	20	β_0	.0037	-.0287	.0361	.0381
		β_1	-.0010	.0271	-.0291	-.0278
50	5	β_0	.0263	-.0816	.1342	.1956
		β_1	-.0128	.0753	-.1010	-.0910
	10	β_0	.0174	-.0389	.0738	.1151
		β_1	-.0088	.0394	-.0571	-.0537
	20	β_0	.0129	-.0141	.0400	.0305
		β_1	.0070	.0174	-.0314	-.0304

Note : All the above result are based on 1000 samples and 500 bootstrap trails in each sample.

Table 4.1 d

Biases of the Estimators Using method 2

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	-.0475	.0921	-.0029	.0169
		β_1	.0641	.1073	.0208	.0323
	10	β_0	-.0306	-.0411	-.0200	.0195
		β_1	.0379	.0557	.0201	.0231
	20	β_0	-.0216	-.0237	-.0194	-.0192
		β_1	.0235	.0314	.0156	.0162
20	5	β_0	.0130	.0276	-.0017	.0059
		β_1	.0027	.0023	.0076	.0119
	10	β_0	.0008	.0129	-.0114	.0013
		β_1	.0053	.0031	.0075	.0082
	20	β_0	-.0008	.0077	-.0093	.0104
		β_1	.0035	.0020	.0051	.0050
30	5	β_0	.0124	.0191	.0057	.0174
		β_1	-.0003	-.0022	.0015	.0006
	10	β_0	.0029	.0103	.0045	-.0061
		β_1	.0022	.0018	.0025	.0031
	20	β_0	.0047	.0104	-.0010	.0003
		β_1	-.0008	-.0014	.0002	-.0004

Table 4.1 d (continued)

Biases of the Estimators Using method 2

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		$\lambda \cdot E$	BOOTSTRAP	N B C	P B C
40	5	β_0	.0070	.0080	.0059	-.0129
		β_1	.0021	.0049	-.0008	.0030
	10	β_0	.0035	.0074	-.0004	.0015
		β_1	.0018	.0030	.0006	.0006
	20	β_0	.0026	.0060	-.0008	-.0006
		β_1	.0001	.0006	-.0004	-.0005
50	5	β_0	.1440	.4541	-.1661	-.0886
		β_1	-.0900	-.1328	-.0471	-.0714
	10	β_0	.0139	.0205	.0073	.0094
		β_1	-.0067	-.0078	-.0055	-.0056
	20	β_0	.0092	.0133	.0050	.0053
		β_1	-.0044	-.0046	-.0041	-.0040

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.1 e

Biases of the Estimators Using method 2

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	-.0301	.1272	.0670	-.0340
		β_1	.0554	.1392	-.0285	.0212
	10	β_0	-.0368	-.0818	.0082	.0094
		β_1	.0436	.0861	.0011	.0038
	20	β_0	-.0109	.0307	.0089	.0083
		β_1	.0166	.0379	-.0047	.0038
20	5	β_0	.0072	.1208	.1353	.0804
		β_1	.0071	.0958	-.0815	.0707
	10	β_0	.0000	-.0606	.0605	.0676
		β_1	.0063	.0517	-.0391	.0363
	20	β_0	-.0001	-.0278	.0276	.0516
		β_1	.0027	.0256	.0202	.0194
30	5	β_0	.0037	.1607	.1681	.2241
		β_1	.0055	.1152	-.1042	.0912
	10	β_0	.0096	.0658	.0850	.0841
		β_1	-.0016	.0521	.0553	.0521
	20	β_0	.0038	.0318	.0394	.0413
		β_1	-.0002	.0271	.0275	.0266

Table 4.1 e (continued)

Biases of the Estimators Using method 2

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.0069	-.1719	.1857	.1550
		β_1	.0031	.1217	-.1154	-.0999
	10	β_0	.0051	-.0795	.0897	.1102
		β_1	-.0001	.0590	-.0591	-.0554
	20	β_0	.0024	-.0380	.0428	.0462
		β_1	-.0004	.0296	-.0305	-.0295
50	5	β_0	.0332	-.1455	.2120	.1171
		β_1	.0167	.1029	-.1362	-.1071
	10	β_0	.0192	.0658	.1042	.1205
		β_1	-.0092	.0505	-.0690	-.0652
	20	β_0	.0167	.0236	.0571	.0597
		β_1	-.0099	.0204	-.0402	-.0392

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.1 f

Biases of the Estimators Using method 2

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	-.0185	.1211	.0810	.0617
		β_1	.0466	.1144	.0212	.0130
	10	β_0	-.0718	.1008	.0653	.0408
		β_1	.0297	.0915	.0320	-.0252
	20	β_0	-.0134	.0562	.0294	.0276
		β_1	.0183	.0522	.0156	.0141
20	5	β_0	.0162	.1408	.1732	.0219
		β_1	.0013	.0795	-.0768	.0274
	10	β_0	.0081	.1242	.1403	.0767
		β_1	.0011	.0838	-.0817	.0708
	20	β_0	-.0004	.0673	.0665	.0471
		β_1	.0034	.0480	.0411	-.0387
30	5	β_0	.0173	.2013	.2359	.0879
		β_1	-.0030	.1118	.1207	.0661
	10	β_0	.0047	.1583	.1677	.0935
		β_1	.0006	.1009	.0997	.0854
	20	β_0	.0067	.0739	.0873	.1020
		β_1	-.0021	.0508	-.0549	-.0518

Table 4.1 f (continue)

Biases of the Estimators Using method 2

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.0099	.2202	.2400	.7330
		β_1	.0011	.1207	-.1180	-.0646
	10	β_0	.0065	-.1703	.1832	.0418
		β_1	.0004	.1078	-.1085	-.0946
	20	β_0	.0037	-.0834	.0908	.0809
		β_1	-.0010	.0554	-.0574	-.0540
50	5	β_0	.0263	.0791	.1317	-.1738
		β_1	-.0128	.0404	-.0661	-.0899
	10	β_0	.0174	-.1683	.2031	.0122
		β_1	-.0088	.1051	-.1228	-.1076
	20	β_0	.0129	-.0786	.1045	.0914
		β_1	.0070	.0527	-.0668	-.0630

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

In the case of no intercept models, whatever the value of n and λ we chose, the m.l.e is the only choice, as rest of the estimators tend to have significant biases.

It can also be seen from Table 4.2 that for increasing values of n and λ , the biases are decreasing, and PBC' is as good as NBC'.

For $n=20$ and $\lambda = 10$ the m.l.e. has negligible bias, but the bootstrap estimator does not seem to provide satisfactory results. However, for large λ , NBC' and PBC' are equally good and seem to improve over the regular bootstrap estimator.

The results in Table 4.2, in the case of no intercept models, show that for increasing values of n and λ , the biases are reducing, but again the m.l.e. seems to perform better over others.

4.2.2 The MSE of the Estimators

In order to compare a biased estimators with different amounts of bias, a useful criterion is the Mean Square Error (MSE) of the estimators, measured from the population value that is being estimated. Formally,

$$\begin{aligned} MSE(\beta) &= E(\beta - \beta)^2 \\ &= V(\beta) + (Bias)^2 \end{aligned} \quad (4.1)$$

which can be estimated by

$$\widehat{MSE}(\beta) = \widehat{V}(\beta) + (\widehat{Bias})^2 \quad (4.2)$$

where $\widehat{MSE}(\beta)$ is an estimate of $MSE(\beta)$ and \widehat{Bias} is an estimate of Bias.

The results in Table 4.3a illustrate that for small values of λ , PBC is not a good estimator. For large values of λ , NBC and PBC are equally precise and have smaller MSE than the m.l.e., for all values of n . For large values of n and λ , the bootstrap estimator may be considered to be equivalent to m.l.e.

From Table 4.3b, we can say that for small values of n and λ , PBC is not good, but as n and λ increase, PBC and NBC almost coincide in terms of MSE. For large n and large λ , the bootstrap estimator is as precise as the m.l.e.

The values in the case of no intercept models indicate that for large n and λ , NBC is more appropriate than the m.l.e. and that the mse of the bootstrap estimator is a little higher than that of m.l.e.

MSE's of different estimators obtained by Method 2 in the Table 4.3d show that for small n , and or λ , the PBC is usually larger than others. NBC is better than the m.l.e. independently of the choice of λ . For large n and λ , PBC and NBC are equally precise and better than the m.l.e.

From Table 4.3e, it is clear that for all values of n and small value of λ , the MSE of PBC is not usually large. However, for all values of n and λ , NBC is better than the m.l.e. and the bootstrap estimator.

In the case of the linear model with no intercept, the MSE of the PBC estimator is unacceptable, as shown in the Table 4.3f when λ is small and n is large. In this case there is no noticeable gain in using another estimate than the m.l.e.

Table 4.2 a

Mean Square Error of the Estimators Using Method 1

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	.7204	.9900	.6074	60.7852
		β_1	.3315	.4614	.2797	.5084
	10	β_0	.3486	.4943	.2432	54.4306
		β_1	.1590	.2324	.1078	.1226
	20	β_0	.1736	.2076	.1447	.1830
		β_1	.0796	.0970	.0651	.0672
20	5	β_0	.4086	.5559	.2979	.3561
		β_1	.1652	.2260	.1183	.1962
	10	β_0	.2097	.2538	.1728	.1774
		β_1	.0853	.1036	.0698	.0886
	20	β_0	.1006	.1097	.0927	.0933
		β_1	.0410	.0447	.0376	.0380
30	5	β_0	.2655	.3430	.2024	.2441
		β_1	.1085	.1110	.0818	.24561
	10	β_0	.1325	.1499	.074	.1172
		β_1	.0547	.0618	.0483	.3410
	20	β_0	.0653	.0689	.0622	.0621
		β_1	.0270	.0285	.0257	.0264

Table 4.2 a (continued)

Mean Square Error of the Estimators Using Method 1

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.1690	.2032	.1397	.1544
		β_1	.0696	.0843	.0570	.0736
	10	β_0	.0843	.0913	.0781	.0779
		β_1	.0347	.0377	.0320	.0324
	20	β_0	.0421	.0437	.0407	.0407
		β_1	.0173	.0180	.0167	.0168
50	5	β_0	.1808	.1528	.1118	16.6205
		β_1	.0843	.0669	.0489	.0501
	10	β_0	.0643	.0690	.0601	.0601
		β_1	.0282	.0302	.0264	.0266
	20	β_0	.0312	.0323	.0303	.0303
		β_1	.0136	.0140	.0132	.0134

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.2 b

Mean Square Error of the Estimators Using Method 1

$$j_0 = 1, \quad j_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	j_0	1.7391	2.5831	1.3415	19.9530
		j_1	.8193	1.255	.6562	.6725
	10	j_0	.8697	1.2393	.6600	86.1792
		j_1	.4134	.6194	.3084	.3193
	20	j_0	.4090	.4873	.3570	.7569
		j_1	.1931	.2380	.1664	.1680
20	5	j_0	1.0070	1.2263	.8387	1.5021
		j_1	.4208	.5292	.3415	.3521
	10	j_0	.4916	.5117	.4481	.4684
		j_1	.2056	.2326	.1853	.1863
	20	j_0	.2435	.2548	.2340	.3211
		j_1	.1024	.1082	.0980	.0981
30	5	j_0	.6434	.7128	.5657	1.4671
		j_1	.2741	.3262	.2376	.2404
	10	j_0	.3099	.3302	.2934	.3947
		j_1	.1327	.1435	.1247	.1250
	20	j_0	.1499	.1544	.1465	.1471
		j_1	.0640	.0665	.0624	.0624

Table 4.2 b (continued)

Mean Square Error of the Estimators Using Method 1

$$f_0 = 1, f_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	f_0	.1075	.4589	.3680	.9031
		f_1	.1735	.2008	.1547	.1558
	10	f_0	.2018	.2133	.1930	.1912
		f_1	.0860	.0922	.0819	.0820
	20	f_0	.0995	.1025	.0974	.0969
		f_1	.0425	.0441	.0415	.0415
50	5	f_0	.3083	.3367	.2866	.3285
		f_1	.1404	.1560	.1299	.1310
	10	f_0	.1499	.1563	.1452	.1444
		f_1	.0675	.0711	.0653	.0653
	20	f_0	.0760	.0776	.0749	.0749
		f_1	.0340	.0348	.0336	.0336

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.2 c

Mean Squares Error of the Estimators Using Method 1

$$\beta_0 = 0, \quad \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	1.0467	1.1580	1.2633	4.9506
		β_1	.5009	.5281	.6847	.7302
	10	β_0	.5191	.7442	.5211	9.3682
		β_1	.2503	.3638	.2727	.2637
	20	β_0	.2526	.4057	.2229	3.8857
		β_1	.1222	.2072	.1131	.1093
20	5	β_0	.5779	.7439	.5238	.7742
		β_1	.2462	.3236	.2386	.2584
	10	β_0	.2898	.3841	.2101	.2999
		β_1	.1234	.1721	.1028	.1043
	20	β_0	.1428	.1722	.1257	.1375
		β_1	.0610	.0768	.0533	.0539
30	5	β_0	.3825	.5758	.2967	.5919
		β_1	.1658	.2669	.1308	.1325
	10	β_0	.1858	.2412	.1517	.2159
		β_1	.0812	.1129	.0674	.0683
	20	β_0	.0894	.1020	.0823	.0959
		β_1	.0392	.0462	.0361	.0362

Table 4.2 c (continued)

Mean Squares Error of the Estimators Using Method 1

$$f_0 = 0, f_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	f_0	.2430	.3342	.2069	.6139
		f_1	.1065	.1538	.0932	.0935
	10	f_0	.1196	.1520	.1044	1.3821
		f_1	.0521	.0700	.0456	.0460
	20	f_0	.0581	.0664	.0551	.0748
		f_1	.0255	.0299	.0240	.0241
50	5	f_0	.1808	.2384	.1628	9.4268
		f_1	.0843	.1157	.0794	.0797
	10	f_0	.0887	.1086	.0805	2.9640
		f_1	.0413	.0529	.0380	.0381
	20	f_0	.0434	.0481	.0417	.0531
		f_1	.0200	.0230	.0194	.0195

Note : All the above results are based on 1000 samples and 500 bootstrap trials
in each sample

Table 4.2 d

Mean Squares Error of the Estimators Using Method 2

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	.7204	.7894	.7119	1.8490
		β_1	.3315	.3675	.3232	1.9075
	10	β_0	.3486	.3903	.3119	.3216
		β_1	.1590	.1795	.1412	.1439
	20	β_0	.1736	.1841	.1639	.1647
		β_1	.0796	.0848	.0748	.0754
20	5	β_0	.4086	.5385	.3323	1.0204
		β_1	.1652	.2190	.1324	.1692
	10	β_0	.2097	.2598	.1673	.2327
		β_1	.0853	.1059	.0678	.0707
	20	β_0	.1006	.1122	.0903	.0963
		β_1	.0410	.0457	.0366	.0371
30	5	β_0	.2655	.3831	.1850	.5407
		β_1	.1085	.1570	.0750	.0886
	10	β_0	.1325	.1674	.1028	.1726
		β_1	.0547	.0695	.0420	.0437
	20	β_0	.0653	.0734	.0581	.0576
		β_1	.0270	.0304	.0239	.0243

Table 4.2 d (continued)

Mean Squares Error of the Estimators Using Method 2

$$I_0 = 1, I_1 = 0$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	I_0	.1690	.2623	.1022	1.5888
		I_1	.0696	.1093	.0409	.0555
	10	I_0	.0843	.1074	.0646	.0649
		I_1	.0347	.0446	.0263	.0274
	20	I_0	.0421	.0476	.0372	.0371
		I_1	.0173	.0197	.0152	.0156
50	5	I_0	.1808	.3117	.1175	.0813
		I_1	.0843	.0558	.0340	.0434
	10	I_0	.0643	.0823	.0492	.0483
		I_1	.0282	.0362	.0215	.0225
	20	I_0	.0312	.0354	.0274	.0274
		I_1	.0136	.0154	.0119	.0121

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.2 e

Mean Squares Error of the Estimators Using Method 2

$$\beta_0 = 1, \quad \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	1.7391	1.9958	1.5367	7.3523
		β_1	.8193	.9357	.7333	.7455
	10	β_0	.8697	.9351	.8148	.8559
		β_1	.4134	.4447	.3887	.3890
	20	β_0	.4090	.4244	.3962	.4023
		β_1	.1931	.2005	.1875	.1874
20	5	β_0	1.0070	1.2860	.8171	1.3322
		β_1	.4208	.5271	.3531	.3661
	10	β_0	.4916	.5651	.4351	.5044
		β_1	.2056	.2352	.1837	.1847
	20	β_0	.2435	.2607	.2297	.5966
		β_1	.1024	.1095	.0971	.0972
30	5	β_0	.6434	.8719	.5129	2.35982
		β_1	.2744	.3626	.2270	.2312
	10	β_0	.3099	.3591	.2781	.6089
		β_1	.1327	.1532	.1202	.1207
	20	β_0	.1499	.1611	.1425	.1442
		β_1	.0640	.0658	.0612	.0613

Table 4.2 e (continued)

Mean Squares Error of the Estimators Using Method 2

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	β_0	.4075	.5780	.3373	5.7766
		β_1	.1735	.2418	.1476	.1558
	10	β_0	.2018	.2400	.1825	.7841
		β_1	.0860	.1020	.0787	.0789
	20	β_0	.0995	.1084	.0949	.0932
		β_1	.0425	.0463	.0409	.0409
50	5	β_0	.3083	.4321	.2730	32.9421
		β_1	.1404	.1912	.1277	.3135
	10	β_0	.1499	.1768	.1409	.1634
		β_1	.0675	.0793	.0642	.0642
	20	β_0	.0760	.0820	.0740	.0736
		β_1	.0340	.0366	.0334	.0334

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.2 f

Mean Squares Error of the Estimators Using Method 2

$$\beta_0 = 0, \quad \beta_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
10	5	β_0	1.0467	1.1111	1.0997	2.6853
		β_1	.5009	.5240	.5737	.6787
	10	β_0	.5191	.6040	.4679	.5260
		β_1	.2503	.2875	.2317	.2369
	20	β_0	.2526	.2777	.2318	.2824
		β_1	.1222	.1338	.1145	.1148
20	5	β_0	.5779	.6851	.6825	7.1194
		β_1	.2462	.2613	.3324	.6092
	10	β_0	.2898	.3863	.2598	1.1101
		β_1	.1234	.1580	.1176	.1243
	20	β_0	.1428	.1718	.1277	.3590
		β_1	.0610	.0723	.0556	.0559
30	5	β_0	.3825	.5266	.5222	93.8736
		β_1	.1658	.2069	.2458	.3230
	10	β_0	.1858	.2804	.1800	3.6885
		β_1	.0812	.1152	.0819	.0874
	20	β_0	.0894	.1127	.0836	.6155
		β_1	.0392	.0486	.0369	.0369

Table 4.2 f (continued)

Mean Squares Error of the Estimators Using Method 2

$$J_0 = 0, J_1 = 1$$

n	λ		M L E	BOOTSTRAP	N B C	P B C
40	5	J_0	.2130	.3606	.3944	421.3743
		J_1	.1065	.1373	.1837	.2564
	10	J_0	.1196	.1969	.1306	9.9268
		J_1	.0521	.0797	.0581	.0590
	20	J_0	.0587	.0781	.0577	.8312
		J_1	.0255	.0332	.0254	.0254
50	5	J_0	.1808	18.4312	18.6564	56.6723
		J_1	.0843	4.4974	4.6159	.2004
	10	J_0	.0887	.1528	.1153	11.6594
		J_1	.0413	.0647	.0534	.0524
	20	J_0	.0131	.0587	.0479	2.3239
		J_1	.0200	.0263	.0221	.0219

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

4.2.3 Estimation of the MSE of Maximum Likelihood Estimators

Method 1 seems to overestimate the MSE, specially for small values of λ , but Method 2 produces satisfactory results in all cases. This may give the basis for hypothesis testing on the parameters. Thus the bootstrap procedure seems to provide valid estimates of the MSE of the m.l.e.'s if we use Method 2.

Table 4.3 a

Comparison of M S E and Est. M S E of $\hat{\beta}$

$$\beta_0 = 1, \quad \beta_1 = 0$$

n	λ		M L E	Method 1	Method 2
10	5	β_0	.7204	2.0833	.7676
		β_1	.3315	1.0991	.3883
	10	β_0	.3486	.6411	.3071
		β_1	.1590	.3010	.1381
	20	β_0	.1736	.1989	.1436
		β_1	.0796	.0936	.0647
20	5	β_0	.4086	.7053	.4781
		β_1	.1652	.2838	.1902
	10	β_0	.2097	.2534	.1963
		β_1	.0853	.1021	.0787
	20	β_0	.1006	.1038	.0930
		β_1	.0110	.0418	.0373
30	5	β_0	.2655	.3833	.3047
		β_1	.1085	.1551	.1238
	10	β_0	.1325	.1458	.1251
		β_1	.0517	.0596	.0513
	20	β_0	.0653	.0647	.0598
		β_1	.0270	.0265	.0246

Table 4.3 a (continued)

Comparison of M S E and Est. M S E of β

$$\beta_0 = 1, \beta_1 = 0$$

n	λ		M L E	Method 1	Method 2
40	5	β_0	.1690	.2253	.2132
		β_1	.0696	.0949	.0879
	10	β_0	.0843	.0958	.0892
		β_1	.0347	.0402	.0373
	20	β_0	.0421	.0448	.0430
		β_1	.0173	.0188	.0180
50	5	β_0	.0980	.1604	.1312
		β_1	.0375	.0695	.0602
	10	β_0	.0643	.0700	.0666
		β_1	.0282	.0304	.0290
	20	β_0	.0312	.0331	.0319
		β_1	.0136	.0144	.0139

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.3 b

Comparison of M S E and Est. M S E of $\hat{\beta}$

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	Method 1	Method 2
10	5	β_0	1.7391	5.1918	1.5952
		β_1	.8193	3.0634	.6967
	10	β_0	.8697	2.2026	.7383
		β_1	.4134	1.2652	.3211
	20	β_0	.4090	.7900	.3592
		β_1	.1931	.4422	.1563
20	5	β_0	1.0070	1.4523	1.1988
		β_1	.4208	.7576	.4705
	10	β_0	.4916	.5971	.5443
		β_1	.2056	.3028	.2120
	20	β_0	.2435	.2696	.2613
		β_1	.1024	.1350	.1020
30	5	β_0	.6434	.8676	.7855
		β_1	.2744	.4566	.3140
	10	β_0	.3099	.3780	.3488
		β_1	.1327	.1955	.1387
	20	β_0	.1499	.1797	.1679
		β_1	.0640	.0924	.0669

Table 4.3 b (continued)

Comparison of M S E and Est. M S E of β

$$\beta_0 = 1, \beta_1 = 1$$

n	λ		M L E	Method 1	Method 2
40	5	β_0	.4075	.6520	.5926
		β_1	.1735	.3467	.2375
	10	β_0	.2018	.2983	.2516
		β_1	.0860	.1569	.1016
	20	β_0	.0995	.1438	.1201
		β_1	.0425	.0753	.0486
50	5	β_0	.3083	.4708	.4460
		β_1	.1404	.2630	.1869
	10	β_0	.1499	.2213	.1932
		β_1	.0675	.1227	.0810
	20	β_0	.0760	.1074	.0894
		β_1	.0340	.0593	.0376

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each sample.

Table 4.3 c

Comparison of M S E and Est. M S E of $\hat{\beta}$

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	Method 1	Method 2
10	5	β_0	1.0467	7.8477	1.2100
		β_1	500 bootstrap	5.5583	.5589
	10	β_0	.5191	3.1732	.5324
		β_1	.2503	2.0339	.2308
	20	β_0	.2526	1.3785	.2445
		β_1	.1222	.8295	.1039
20	5	β_0	.5779	1.7695	2.0863
		β_1	.2462	1.0490	.6801
	10	β_0	.2898	.7944	.4761
		β_1	.1234	.4512	.1873
	20	β_0	.1428	.3292	.1981
		β_1	.0610	.1821	.0762
30	5	β_0	.3825	1.7280	1.9721
		β_1	.1658	1.1129	.5861
	10	β_0	.1858	.5099	.3737
		β_1	.0812	.2927	.1479
	20	β_0	.0894	.2099	.1317
		β_1	.0392	.1186	.0510

Table 4.3 c (continued)

Comparison of M S E and Est. M S E of β

$$\beta_0 = 0, \beta_1 = 1$$

n	λ		M L E	Method 1	Method 2
40	5	β_0	.2430	.9104	1.9498
		β_1	.1065	.5372	.6326
	10	β_0	.1196	.3997	.3096
		β_1	.0521	.2303	.1217
	20	β_0	.0587	.1762	.0985
		β_1	.0255	.1004	.0389
50	5	β_0	.1808	.6889	9148.25
		β_1	.0843	.4301	2214.80
	10	β_0	.0887	.3061	.2652
		β_1	.0413	.1866	.1108
	20	β_0	.0434	.1384	.0814
		β_1	.0200	.0836	.0334

Note : All the above results are based on 1000 samples and 500 bootstrap trials in each samples.

Bibliography

- [1] Babu, G.J. (1992) Personal Communication.
- [2] Bhattacharya, G.K., and Fries, A. (1982) Inverse Gaussian Regression and Accelerated Life Tests, in *Proceedings of the special Topics Meeting on Survival Analysis*, 138 th Meeting of the Institute of Mathematical Statistics, Columbus, Ohio, Oct. 1981, 101-117.
- [3] Bhattacharya, G.K., and Fries, A. (1986) On the Inverse Gaussian Multiple Regression and Model Checking Procedures, *Reliability and Quality Control*, Tech. Rep., Univ. of Wisconsin, 87-100.
- [4] Efron, B. (1982) The jackknife, the Bootstrap and other resampling Plans, *SIAM, Philadelphia*.
- [5] Freedman, D. and Peters, S. (1981) Bootstrapping regression models, *Ann. Statist.*, **9**, 1218-1228.
- [6] Gray, H.L., Schucany, W.R. (1972) *The Generalized Jackknife statistic*, New York, M. Dekker, 303-306.

- [7] Hinkley, D. (1977) On jackknifing in unbalanced situations, *Technometrics*, **19**, 285-292.
- [8] Miller, R.G. (1974b) An unbalanced jackknife, *Ann. statist.*, **2**, 880-891.
- [9] Quenouille, M.H. (1956) Notes on bias in estimation, *Biometrika*, **43**, 353-360.
- [10] Rao, C.R. (1973) *Linear Statistical Inference and its Applications*, New York, Wiley.
- [11] Schrodinger, E. (1915) Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung, *Physikalische Zeitschrift*, **16**, 289-295.
- [12] Tukey, J.W. (1958) Bias and confidence in not quite large samples (Abstract), *Ann. statist.*, **29**, 611.
- [13] Tweedie, M. C. K. (1957) Statistical properties of inverse gaussian distribution, *Ann. Math. Statist.*, **28**, 362-377.
- [14] Wald, A. (1947) *Sequential Analysis*, New York, Wiley.
- [15] White, H. (1984) *Asymptotic Theory for Econometricians*, New York, Academic Press.
- [16] Whitmore, G. A. (1980) An Exact Multiple Regression Model for Inverse Gaussian Data, working paper, McGill University, Faculty of Management.

Appendix : Derivation of (3.22)

For the IG regression model as described in section 3.1.1, the estimator of β is given by

$$\beta = (X'YX)^{-1}X'1 = S^{-1}X'1 \quad (A.1)$$

where $S = X'YX$ which further can be written as

$$S = X'MX + X'(Y - M)X$$

$$(I + (X'RX)F^{-1})F$$

where $R = Y - M$ and $F = X'MX$. Therefore,

$$S^{-1} = F^{-1}(I + (X'RX)F^{-1})^{-1}$$

$$F^{-1}(I + A)^{-1}$$

$$F^{-1} \sum_{j=0}^{\infty} (-1)^j \frac{A^j}{j!} \quad (A.2)$$

where $A = X'RXF^{-1}$ and

$$E(S^{-1}) = E(F^{-1} \sum_{j=0}^{\infty} \frac{A^j}{j!})$$

$$E(F^{-1}) = F^{-1} (1 + F^{-1} \frac{A^2}{2!} + F^{-1} \frac{A^3}{3!} + \dots)$$

$$\begin{aligned}
&= F^{-1} + F^{-1}E(A) + \frac{1}{2}E(A^2) + \frac{1}{6}F^{-1}E(A^3) + \dots \\
&= (X'MX)^{-1} + \frac{1}{2}F^{-1}E(X'RXF^{-1}X'RX)F^{-1}
\end{aligned}$$

which implies that

$$E(\hat{\beta}) = F^{-1}X'1 + \frac{1}{2}F^{-1}E(X'RXF^{-1}X'RX)F^{-1}X'1 \quad (A.3)$$

Now consider

$$\begin{aligned}
E(X'RXF^{-1}X'RX) &= E\left\{\left(\sum_{i=1}^n R_i X_i X_i'\right)F^{-1}\left(\sum_{i=1}^n R_i X_i X_i'\right)\right\} \\
&= E\left\{\sum_{i=1}^n R_i^2 X_i X_i' F^{-1} X_i X_i' + \sum_{i \neq j} R_i R_j X_i X_i' F^{-1} X_j X_j'\right\} \\
&= \sum_i V(R_i) X_i X_i' F^{-1} X_i X_i' + \sum_{i \neq j} Cov(Y_i, Y_j) X_i X_i' F^{-1} X_j X_j' \\
&\quad + \sum_{i=1}^n \sigma_i^2 X_i X_i' F^{-1} X_i X_i' \quad (A.4)
\end{aligned}$$

where $E(R_i) = 0$, $V(R_i) = \frac{\mu_i^2}{N}$ and $Cov(Y_i, Y_j) = 0$. Also $F^{-1} = (X'MX)^{-1}$

$\sum_i \mu_i X_i X_i'$ is of $O(n)$ and hence, F^{-1} is of $O(1/n)$ and

$$\begin{aligned}
E(S^{-1}X'1) &= F^{-1}X'1 + \frac{1}{2}F^{-1}\sum_{i=1}^n \sigma_i^2 X_i X_i' F^{-1} X_i X_i' F^{-1}X'1 \\
&= F^{-1}X'1 + \frac{1}{2}F^{-1}\sum_{i=1}^n \sigma_i^2 X_i X_i' F^{-1} X_i X_i' \beta \\
&\quad \beta + O(1/n)
\end{aligned}$$

Hence

$$E(\beta) = \beta + O(1/n) \quad (A.5)$$

We can say that the bias of β is $\frac{1}{2}F^{-1}\sum \sigma_i^2 X_i X_i' F^{-1} X_i X_i' F^{-1} X_i X_i' \beta$ which is of order $O(1/n)$ under the condition that $(X'YX)^{-1}$ is of order $O(1/n)$.